UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/785,238	02/24/2004	Takeshi Otani	FUJR 20.949	1025

26304 7590 01/31/2008
KATTEN MUCHIN ROSENMAN LLP
575 MADISON AVENUE
NEW YORK, NY 10022-2585

EXAMINER	
COLUCCI, MICHAEL C	

ART UNIT	PAPER NUMBER
2626	

MAIL DATE	DELIVERY MODE
01/31/2008	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)	
	10/785,238	OTANI ET AL.	
	Examiner	Art Unit	
	Michael C. Colucci	2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☐ Responsive to communication(s) filed on _____.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-2 and 28-30* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-2 and 28-30* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *24 February 2004* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☒ Some * c)☐ None of:

1.☒ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☒ Information Disclosure Statement(s) (PTO/SB/08)
Paper No(s)/Mail Date *See Continuation Sheet*.

4) ☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____ .

5) ☐ Notice of Informal Patent Application

6) ☐ Other: _____.

Continuation of Attachment(s) 3). Information Disclosure Statement(s) (PTO/SB/08), Paper No(s)/Mail Date :02/01/2007, 03/23/2004, 03/17/2004.

DETAILED ACTION

Response to Arguments

1. Applicant's arguments, see, filed 11/19/2007, with respect to the rejection(s) of claim(s) 1 and 2 under 102(b) have been fully considered and are persuasive. Therefore, the rejection has been withdrawn. However, upon further consideration, a new ground(s) of rejection is made in view of Ananthaiyer et al US 6385548 B2 (hereinafter Ananthaiyer). The prior art of Borth and Ananthaiyer both disclose threshold comparison as a means for testing whether magnitudes are below a threshold which correspond to noise. When magnitude levels are below one or more threshold levels, noise is said to be present and a new noise can be estimated to produce a more suppressed/flat signal. Likewise, when a second threshold is breached, a new estimate for voice/speech can be produced which would clarify an input voice segment. Additionally, the normalization of factors such as a smoothing factor is used to smooth background noise per frame relative to the old and new estimated noise produced in dependency on a threshold. However, summing differences and using averages for the division on a per frame basis is not particularly taught by Borth. Additionally, the motivational aspect to combine the previous prior arts of Tsutsui RE36683 and O'Hagan US 5581658 would not produce a combined reference to reject the limitation for summing differences and dividing by averages for a purpose of normalization on a per frame basis.

Claim Rejections - 35 USC § 101

2. 35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

The claimed invention is directed to non-statutory subject matter. Claims 1, 2, and 28-30 disclose a process of determining talk spurts in an "input signal", where an input can by any type of signal. Therefore, claims 1, 2, and 28-30 are construed to be a discovery of a mathematical algorithm where data such as a talk spurt is merely extracted from a general input signal. However, the claimed operation would be potentially statutory if the claimed "input signal" were a specific type of signal (i.e. input voice/noise signal, input speech signal, input music signal, input conversation signal, etc.).

Claim Rejections - 35 USC § 103

3. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

4. Claims 1, 2, and 28 rejected under 35 U.S.C. 103(a) as being unpatentable over Borth et al, US 4630304 (hereinafter Borth) in view of Ananthaiyer et al US 6385548 B2 (hereinafter Ananthaiyer).

Re claims 1 and 28, Borth teaches a voice activity detector that detects talk spurts in an input signal (col 2 line 46 – col 3 line 6), comprising:

a frequency spectrum calculator that calculates a frequency spectrum of the input signal (col 3 line 53 – col 4 line 9);

a flatness evaluator that finds a maximum value of the frequency spectrum (col 5 lines 17-31),

NOTE: finding a maximum of a frequency spectrum as portrayed in figure 15 of the drawings is construed to be functionally equivalent as a threshold that is used to distinguish noise from speech in a signal (col 5 lines 17-31), where the max value would be construed to be a value just below the threshold level that can be considered noise.

a voice/noise discriminator that determines whether the input signal contains a talk spurt (col 2 line 46 – col 3 line 6), by comparing the normalized flatness factor of the frequency spectrum with a predetermined threshold (col 5 lines 17-31).

However, Borth fails to teach adding up differences between spectral components and the maximum value thereof; and generates the resulting sum of the differences as a flatness factor indicating flatness of the frequency spectrum, to teach wherein said flatness evaluator calculates an average of spectral components of the input signal normalizes the resulting sum of the differences by dividing by the calculated average and outputs a normalized flatness factor (Ananthaiyer col 8 lines 32-55);

Ananthaiyer teaches a process to determine whether a signal frame/portion is noise, tone, or voice by determining an Average Magnitude Difference Function (AMDF)

value 652 for each of a predetermined range of pitch frequencies K over the intervals; second logic 656 for determining an average difference AMDF value over the intervals equal to the sum of the difference between a first minimum AMDF value from each interval m and a second minimum AMDF value from each interval (m-1); sixth logic 664 for computing a second metric equal to the average difference AMDF value over the intervals divided by the sum of the AMDF values over the intervals; and seventh logic 666 for utilizing said first metric and said second metric to determine whether the signal is one of a noise signal, a tone signal, and a voice signal. Particularly, the difference between adjacent/consecutive spectral values is taught where the difference of a spectral value at (m) and (m-1) is computed.

Ananthaiyer also teaches logic that determines if the signal is a noise signal in step 404. In step 404, the signal is characterized as noise, and the logic proceeds to step 410, if any of a number of conditions is true. First, the signal is characterized as noise if the AMDF.sub.sum is equal to zero. This case represents the detection of absolute silence. Second, the signal is characterized as noise if the AMDF.sub.norm for the current detection cycle i is greater than a threshold N, representing a large value of AMDF.sub.norm. Finally, the signal is characterized as noise if the signal detected in the previous detection cycle (i-1) was noise and the AMDF.sub.norm is greater than a threshold N2N which is less stringent than N. This condition applies the rule from the first observed characteristic described above, specifically that the threshold for detecting subsequent noise signals can be made less stringent (Fig. 4).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention summing the differences of between consecutive magnitude components and producing a normalized flatness factor by dividing the sum by the average of the spectral components. Using the sum of differences diving by the average would allow for the conservation of processing resources, where noise and speech can be distinguished in segments rather than globally, which would require more processing time and memory. Operations of calculating a moving average (frame by frame average) would allow for a smoothed signal that can be weighted for the purpose of suppressing unwanted noise in a signal. Additionally, by computing the difference prior to dividing by the average, a division by zero or a negative number can be avoided on a per frame basis, which would allow for more accurate distinguishing between noise and sound in addition to a faster process overall.

Re claim 2, Borth teaches the voice activity detector according to claim 1, wherein: the input signal is provided on a frame basis (col 1 line 50 – col 2 line 9); and said frequency spectrum calculator comprises either a spectral analyzer that analyzes the given signal frame in frequency domain, or a plurality of band pass filters (col 4 lines 10-26) that divide the given signal frame into individual frequency components so as to calculate power of each frequency component (col 2 line 46 – col 3 line 6).

5. Claim 29 rejected under 35 U.S.C. 103(a) as being unpatentable over Borth et al, US 4630304 (hereinafter Borth) in view of Ananthaiyer et al US 6385548 B2

(hereinafter Ananthaiyer) and further in view of Sugar et al USPGPUB 20030198304 A1 (hereinafter Sugar).

Re claim 29, Borth teaches a voice activity detector that detects talk spurts in an input signal (col 2 line 46 – col 3 line 6), comprising:

a frequency spectrum calculator that calculates a frequency spectrum of the input signal (col 3 line 53 – col 4 line 9);

a flatness evaluator that calculates a flatness factor FLT indicating flatness of the frequency spectrum (col 5 lines 17-31);

NOTE: finding a maximum of a frequency spectrum as portrayed in figure 15 of the drawings is construed to be functionally equivalent as a threshold that is used to distinguish noise from speech in a signal (col 5 lines 17-31), where the max value would be construed to be a value just below the threshold level that can be considered noise.

a voice/noise discriminator that determines whether the input signal contains a talk spurt, by comparing the flatness factor FLT of the frequency spectrum with a predetermined threshold THR (col 5 lines 17-31),

However, Borth fails to teach wherein said flatness evaluator calculates an average Pm of spectral components of the input signal (Ananthaiyer col 8 lines 32-55), determines the threshold THR from the average Pm (Ananthaiyer col 8 lines 32-55)

Ananthaiyer teaches a process to determine whether a signal frame/portion is noise, tone, or voice by determining an Average Magnitude Difference Function (AMDF) value 652 for each of a predetermined range of pitch frequencies K over the intervals;

second logic 656 for determining an average difference AMDF value over the intervals equal to the sum of the difference between a first minimum AMDF value from each interval m and a second minimum AMDF value from each interval (m-1); sixth logic 664 for computing a second metric equal to the average difference AMDF value over the intervals divided by the sum of the AMDF values over the intervals; and seventh logic 666 for utilizing said first metric and said second metric to determine whether the signal is one of a noise signal, a tone signal, and a voice signal. Particularly, the difference between adjacent/consecutive spectral values is taught where the difference of a spectral value at (m) and (m-1) is computed.

Ananthaiyer also teaches logic that determines if the signal is a noise signal in step 404. In step 404, the signal is characterized as noise, and the logic proceeds to step 410, if any of a number of conditions is true. First, the signal is characterized as noise if the AMDF.sub.sum is equal to zero. This case represents the detection of absolute silence. Second, the signal is characterized as noise if the AMDF.sub.norm for the current detection cycle i is greater than a threshold N, representing a large value of AMDF.sub.norm. Finally, the signal is characterized as noise if the signal detected in the previous detection cycle (i-1) was noise and the AMDF.sub.norm is greater than a threshold N2N which is less stringent than N. This condition applies the rule from the first observed characteristic described above, specifically that the threshold for detecting subsequent noise signals can be made less stringent (Fig. 4).

However, Borth in view of Ananthaiyer fails to teach counting a number of spectral components that exceed the threshold THR, and uses the resulting number as the flatness factor FLT of the frequency spectrum (Sugar [0231]),

the flatness factor FLT and the threshold THR are given by

FLT = count (P[k] > THR) (Sugar [0231])

THR = Pm * COEFF (Sugar Fig. 19 items 450 and 452)

where L and M are lower and upper ends of a frequency range of the input signal, k is a frequency, P[k] is a spectral component of frequency k, count() is the number of the spectral components that exceed the threshold THR (Sugar [0231]), and COEFF is a multiplication factor (Sugar Fig. 19 items 450 and 452).

Sugar teaches computing Fast Fourier Transform (FFT) values at a plurality of frequency bins from a digital signal representing activity in a frequency band during a time interval; computing the power at each frequency bin; adding the power at each frequency bin for a current time interval with the power at the corresponding frequency bin for a previous time interval to obtain a running sum of the power at each frequency bin; comparing the power at each frequency bin with a power threshold to obtain a duty count of the number of times that the power at each frequency bin exceeds the power threshold over time intervals; and comparing the power at each frequency bin for a current time interval with the power at the corresponding frequency bin for a previous time interval to track the maximum power in each frequency bin over time intervals. This process may also be implemented by instructions encoded on a processor

readable medium that, when executed by a processor, cause the processor to perform these same steps.

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention distinguishing between noise and voice by use of thresholds to determine how flat/suppressed a signal is, where the number of times a threshold is crossed is counted relative to an average value in the frequency spectrum. Finding magnitudes relative to an average on a frame by frame basis would allow for conservation of processing resources, where noise and speech can be distinguished in segments rather than globally, which would require more processing time and memory. Additionally, by computing the difference prior to dividing by the average, a division by zero or a negative number can be avoided on a per frame basis, which would allow for more accurate distinguishing between noise and sound in addition to a faster process overall. Operations of calculating a moving average (frame by frame average) would allow for a smoothed signal that can be weighted for the purpose of suppressing unwanted noise in a signal. Additionally, counting threshold values that are crossed would allow for additional conservation of resources, where a count would be the only data present relative to the number of times that a maximum value that is present. As a result of counting particular magnitudes, a faster comparison of data from previous sets can be accomplished as there would more resources that can be dedicated to comparing a count value rather than comparing all magnitudes in a frame.

6. Claim 30 rejected under 35 U.S.C. 103(a) as being unpatentable over Borth et al, US 4630304 (hereinafter Borth) in view of Sugar et al USPGPUB 20030198304 A1 (hereinafter Sugar).

Re claim 30, Borth teaches a voice activity detector that detects talk spurts in an input signal (col 2 line 46 – col 3 line 6), comprising:

a frequency spectrum calculator that calculates a frequency spectrum of the input signal (col 3 line 53 – col 4 line 9);

a flatness evaluator that calculates a flatness factor FLT indicating flatness of the frequency spectrum (col 5 lines 17-31);

NOTE: finding a maximum of a frequency spectrum as portrayed in figure 15 of the drawings is construed to be functionally equivalent as a threshold that is used to distinguish noise from speech in a signal (col 5 lines 17-31), where the max value would be construed to be a value just below the threshold level that can be considered noise.

a voice/noise discriminator that determines whether the input signal contains a talk spurt, by comparing the flatness factor FLT of the frequency spectrum with a predetermined threshold THR (col 5 lines 17-31),

wherein said flatness evaluator finds a maximum value PMAX of the frequency spectrum, determines the threshold THR from the maximum value PMAX, counts a number of spectral components that exceed the threshold THR (Sugar [0231]), and uses the resulting number as the flatness factor FLT of the frequency spectrum,

the flatness factor FLT and the threshold THR are given by

FLT = count (P[k] > THR) (Sugar [0231])

THR = Pm * COEFF (Sugar Fig. 19 items 450 and 452)

where L and M are lower and upper ends of a frequency range of the input signal, k is a frequency, P[k] is a spectral component of frequency k, count() is the number of the spectral components that exceed the threshold THR (Sugar [0231]), and COEFF is a multiplication factor (Sugar Fig. 19 items 450 and 452).

Sugar teaches computing Fast Fourier Transform (FFT) values at a plurality of frequency bins from a digital signal representing activity in a frequency band during a time interval; computing the power at each frequency bin; adding the power at each frequency bin for a current time interval with the power at the corresponding frequency bin for a previous time interval to obtain a running sum of the power at each frequency bin; comparing the power at each frequency bin with a power threshold to obtain a duty count of the number of times that the power at each frequency bin exceeds the power threshold over time intervals; and comparing the power at each frequency bin for a current time interval with the power at the corresponding frequency bin for a previous time interval to track the maximum power in each frequency bin over time intervals. This process may also be implemented by instructions encoded on a processor readable medium that, when executed by a processor, cause the processor to perform these same steps.

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention distinguishing between noise and voice by use of thresholds to determine how flat/suppressed a signal is, where the number of times a threshold is

crossed is counted relative to a maximum value in the frequency spectrum. Finding magnitudes relative to an maximum vale on a frame by frame basis would allow for conservation of processing resources, where noise and speech can be distinguished in segments rather than globally, which would require more processing time and memory. Additionally, by computing the difference prior to dividing by the average, a division by zero or a negative number can be avoided on a per frame basis, which would allow for more accurate distinguishing between noise and sound in addition to a faster process overall. Additionally, counting threshold values that are crossed would allow for additional conservation of resources, where a count would be the only data present relative to the number of times that a maximum value that is present. As a result of counting particular magnitudes, a faster comparison of data from previous sets can be accomplished as there would more resources that can be dedicated to comparing a count value rather than comparing all magnitudes in a frame.

Conclusion

7. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. US 5774849 A, US 5581658 A, US RE36683, US 6862567 B1, US 5666466 A, US 5920834 A, US 5717724 A, US 5537509 A, US 5475712 A, US 5536902 A, US 5479522 A, US 6144937 A, US 4630305 A, US 6084967 A, US 20020188445 A1, US 7031916 B2, US 6999520 B2, US 5189701 A.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Michael C. Colucci whose telephone number is (571)-270-1847. The examiner can normally be reached on 9:30 am - 6:00 pm, Monday-Friday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on (571)-272-7602. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

RICHEMOND DORVIL
SUPERVISORY PATENT EXAMINER

Michael Colucci Jr.
Patent Examiner
AU 2626
(571)-270-1847
Michael.Colucci@uspto.gov